

VU Research Portal

Concurrent validity and interrater reliability of a new smartphone application to assess 3D active cervical range of motion in patients with neck pain

Stenneberg, Martijn S.; Busstra, Harm; Eskes, Michel; van Trijffel, Emiel; Cattrysse, Erik; Scholten-Peeters, Gwendolijne G.M.; de Bie, Rob A.

published in

Musculoskeletal science & practice
2018

DOI (link to publisher)

[10.1016/j.msksp.2017.12.006](https://doi.org/10.1016/j.msksp.2017.12.006)

document version

Publisher's PDF, also known as Version of record

document license

Article 25fa Dutch Copyright Act

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Stenneberg, M. S., Busstra, H., Eskes, M., van Trijffel, E., Cattrysse, E., Scholten-Peeters, G. G. M., & de Bie, R. A. (2018). Concurrent validity and interrater reliability of a new smartphone application to assess 3D active cervical range of motion in patients with neck pain. *Musculoskeletal science & practice*, 34(April), 59-65. <https://doi.org/10.1016/j.msksp.2017.12.006>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl



Original article

Concurrent validity and interrater reliability of a new smartphone application to assess 3D active cervical range of motion in patients with neck pain



Martijn S. Stenneberg^{a,b,*}, Harm Busstra^{a,c}, Michel Eskes^a, Emiel van Trijffel^{a,c}, Erik Cattrysse^c, Gwendolijne G.M. Scholten-Peeters^{a,d}, Rob A. de Bie^{a,b,e}

^a SOMT University of Physiotherapy, Amersfoort, The Netherlands

^b CAPHRI School for Public Health and Primary Care, Maastricht University, The Netherlands

^c Vrije Universiteit Brussel, Department of Experimental Anatomy, Brussels, Belgium

^d Department of Human Movement Sciences, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam Movement Sciences, The Netherlands

^e Maastricht University, Department of Epidemiology, Maastricht, The Netherlands

ARTICLE INFO

Keywords:

Active range of motion

Assessment

Nonspecific neck pain

Reliability and validity

Smartphone

ABSTRACT

Background: There is a lack of valid, reliable, and feasible instruments for measuring planar active cervical range of motion (aCROM) and associated 3D coupling motions in patients with neck pain. Smartphones have advanced sensors and appear to be suitable for these measurements.

Objectives: To estimate the concurrent validity and interrater reliability of a new iPhone application for assessing planar aCROM and associated 3D coupling motions in patients with neck pain, using an electromagnetic tracking device as a reference test.

Design: Cross-sectional study.

Methods: Two samples of neck pain patients were recruited; 30 patients for the validity study and 26 patients for the reliability study. Validity was estimated using intraclass correlation coefficients (ICCs), and by calculating 95% limits of agreement (LoA). To estimate interrater reliability, ICCs were calculated. Cervical 3D coupling motions were analyzed by calculating the cross-correlation coefficients and ratio between the main motions and coupled motions for both instruments.

Results: ICCs for concurrent validity and interrater reliability ranged from 0.90 to 0.99. The width of the 95% LoA ranged from about 5° for right lateral bending to 11° for total rotation.

No significant differences were found between both devices for associated coupling motion analysis.

Conclusions: The iPhone application appears to be a useful discriminative tool for the measurement of planar aCROM and associated coupling motions in patients with neck pain. It fulfills the need for a valid, reliable, and feasible instrument in clinical practice and research. Therapists and researchers should consider measurement error when interpreting scores.

1. Introduction

Assessment of active cervical range of motion (aCROM) is an important part of the clinical examination of impairments in patients with neck pain (Childs et al., 2008; de Koning et al., 2008; Williams et al., 2010). Nearly all patients with neck pain show a decreased aCROM compared to healthy subjects (Stenneberg et al., 2017). Current physical therapy guidelines recommend to assess and restore cervical range of motion in order to reduce neck complaints (Bier et al., 2016; Childs et al., 2008).

It is essential for physical therapists to use valid measurement instruments with high levels of interrater reliability when obtaining quantitative data on the amount of aCROM in patients (Kottner et al., 2011; Streiner and Norman, 2008; de Vet et al., 2006). Instruments such as visual estimation, single inclinometers, and the cervical range of motion device (CROM-device), of which the latter two possess adequate measurement properties, are available to measure aCROM (Antonaci et al., 2000; Jordan, 2000; de Koning et al., 2008; Williams et al., 2010). All of these instruments, however, only provide measurements of primary planar movements, which may not be sufficient to register

* Corresponding author. SOMT, SOMT University of Physiotherapy, Department of Manual Therapy, Softwareweg 5, 3821 BN, Amersfoort, The Netherlands.
E-mail address: m.stenneberg@somtuniversity.nl (M.S. Stenneberg).

intricate cervical movements (Jordan, 2000).

Cervical kinematics are complex (Antonaci et al., 2000; Bogduk and Mercer, 2000; Cook et al., 2006; Feipel et al., 1999) and are likely to change as a consequence of neck pain (Guo et al., 2012; Woodhouse and Vasseljen, 2008). Not only is planar range of motion of the primary movements (flexion-extension, lateral bending and rotation) reduced in neck pain patients compared to healthy controls, but also 3D coupling between cervical rotation, lateral bending, and flexion-extension is affected as well (Guo et al., 2012; Woodhouse and Vasseljen, 2008). To evaluate these 3D coupling motions, instruments should be capable of measuring these adequately (Guo et al., 2012).

To measure both the planar movements of the cervical spine and the associated coupled motions, devices such as electromagnetic tracking systems and optical motion systems are available and accepted as a gold standard (Antonaci et al., 2000; Nafis et al., 2006; Williams et al., 2010). These systems, however, are expensive, difficult to use, and the interpretability of data is complex. Therefore, these are not ideally suited for use in clinical practice.

To date, there is a growing tendency to use smartphone applications for assessment of range of motion in different joints (Ferriero et al., 2011; Kolber et al., 2013; Ockendon and Gilbert, 2012; Shin et al., 2012). Current smartphones have high-quality built-in microelectromechanical systems (MEMS) such as gyroscopes and accelerometers. Using integrated gyroscopic data allows assessment of 3D neck orientation (Jasiewicz et al., 2007; Luinge, 1999; Theobald et al., 2012).

Smartphone applications for measuring planar aCROM are clinically applicable and have moderate to good measurement properties (Guidetti et al., 2016; Kolber et al., 2013; Quek et al., 2014; Tousignant-Lafamme et al., 2013). However, all these studies included healthy participants, which limits generalizability of their results. The study sample should reflect the population of interest because reliability is highly dependent on the distribution of aCROM in the population (de Vet et al., 2011).

Moreover, the reliability and validity of the reference tests, i.e., goniometers and inclinometers, which are used in most of these studies are questionable, therewith biasing claimed measurement properties of the applications (Guidetti et al., 2016; Kolber et al., 2013; Tousignant-Lafamme et al., 2013). Furthermore, to the best of our knowledge, none of the existing applications are able to conduct analysis of the 3D coupling motions of the cervical spine.

Due to lack of valid, reliable, and feasible instruments for measuring planar aCROM and associated coupling motions, we developed a new iPhone-application. The aim of this study was to estimate the concurrent validity and interrater reliability of this new iPhone-application to assess aCROM in patients with nonspecific neck pain presenting in primary care physical therapy, using an electromagnetic tracking device as a reference test.

2. Materials and methods

2.1. Design

A cross-sectional study in which concurrent validity and interrater reliability of the Apple iPhone 4s was estimated using two samples of neck pain patients, with one sample being used for the validity study and the other for the reliability study. An electromagnetic tracking device (Polhemus Liberty) was used as a reference test.

2.2. Participants

Patients from both samples were recruited from five primary care physical therapy practices in the Netherlands. Patients were considered eligible for inclusion when above 18 years, with neck pain existing for at least one week and defined as grade I or II by the Neck Pain Task Force (Guzman et al., 2009). Participants had to be able to read the Dutch language adequately in order to be able to complete the

questionnaires. Patients with previous surgery of the cervical spine or suffering from neurological, visual or vestibular disorders, cervical radiculopathy, or dizziness were excluded.

All participants signed an informed consent. The Central Committee on Research Involving Human Subjects (CCMO, The Hague, The Netherlands) approved the study protocol (non-WMO statement).

2.3. Raters

Two manual therapists (HB, MS) with more than 10 years of clinical experience performed all measurements according to a rigidly standardized protocol. Both raters attended two 4-h training sessions to become familiar with use of the measurement devices and standardization of the examination procedures.

2.4. Self-reported measures

All patients completed a questionnaire to register demographics (age, sex, height and weight, and duration of symptoms) and scored their pain intensity on a 10-point Numeric Pain Rating Scale (NPRS) and completed the Neck Disability Index (NDI) (Vernon and Mior, 1991). The NPRS and NDI are widely used and reliable and valid measurement tools to assess pain and disability in individuals with neck pain (MacDermid et al., 2009; Pool et al., 2007; Schellingerhout et al., 2012).

2.5. Instruments

2.5.1. iPhone application

An Apple iPhone 4s device with the operation system IOS 7.0 (Apple Inc. Cupertino, California, United States) was used. To measure 3D aCROM, we developed an application called '3D range of motion'. This application uses the built-in gyroscope and accelerometer function of the smartphone as angular measurement tools. A CMMotionManager class developed by Apple Inc. converts raw and processed accelerometer, gyroscope, and magnetometer data into range of motion along three axes: x, y, and z axis. The iPhone application records data at a sampling rate of 50 Hz.

2.5.2. Polhemus Liberty

The Polhemus Liberty (Polhemus, Colchester, Vermont, United States), a three dimensional electromagnetic tracking device, was used as a reference test. The Polhemus Liberty is a valid and reliable instrument for measuring spinal range of motion when compared to a Vicon motion capture system (Kaliarantas et al., 2009). The Polhemus Liberty detects angular motions with an accuracy of 0.3° (Nafis et al., 2006). Tracking data were obtained at a sampling rate of 240 Hz. The data of the Polhemus Liberty were processed using Mathcad software (version 14).

2.6. Measurement procedures

2.6.1. Preliminary validation

Prior to the main study, preliminary validation tests were carried out to investigate to which extent the iPhone displays the same angle as the Polhemus Liberty. For this purpose, a wooden phantom (frame) was used with a lever arm that could move left and right to 75°, with 15° stages (Fig. 1). The transmitter of the Polhemus Liberty was placed horizontally, squared, and centered. The measurement devices were placed on the phantom, with 25 cm interspace to avoid interference between the two devices. Before each test, both measurement devices were calibrated in the zero-position. One of the researchers (MS) performed the movements with the arm of the phantom, while simultaneously capturing data from the two devices.

The arm of the phantom was moved into respectively 15°, 30°, 45°, 60°, and 75° and then back to the starting point, in five different

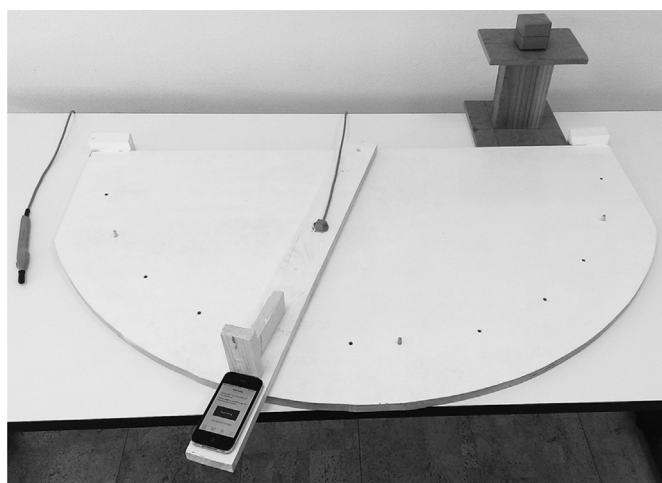


Fig. 1. Preliminary validation study setup and specific positions of the equipment.



Fig. 2. Positioning of the iPhone and sensor of the Polhemus Liberty.

measurement cycles. For each angle, three trials were carried out in succession. Movements to the left and right were recorded separately. This test procedure was repeated for each of the three axes (x , y , and z axis), by changing the orientation of the iPhone in the concerning position.

Mean differences between the Polhemus Liberty and the iPhone were found to be very small, ranging from 0.02° (95%CI -0.15° – 0.19°) for flexion-extension trial 2, to 0.75° (95%CI 0.56° – 0.94°) for rotation trial 3 (Table 1). Because these differences were substantially smaller than those of the instruments currently used in practice (Fletcher and Bandy, 2008; Law and Chiu, 2013), we considered our results sufficient to proceed with the main study.

2.6.2. Measurement procedures main study

ACROM was assessed with participants sitting on a plastic chair with a backrest and a seat height of 50 cm. Their feet were resting on the ground, knees and hips flexed 90° , hands on thighs, and with neutral head-neck position. The iPhone was fastened securely on the forehead by means of an iPhone holder with a rigid Velcro strap (Fig. 2).

Prior to the assessment, there was an introduction to the procedure and participants performed all movements to become familiar with the measurements. Participants were asked to move their head three times, fluently, and as far as possible without provoking pain, and avoiding any movements of the shoulder girdle and thorax. The raters supervised to ensure that the movements were executed properly and followed the relevant axis.

Participants performed a full cycle flexion-extension, left and right rotation, and lateral bending, each three times and with a 10-s interval between each cycle.

To eliminate a potential influence of a fixed movement sequence on the aCROM measurements, the movement sequence (flexion-extension, lateral bending, rotation) was randomized per participant.

Data were recorded directly on the iPhone and sent to the outcome assessor after the measurements were performed. Raters and patients were blinded to their own and each others' results.

2.6.3. Concurrent validity

To avoid interference between the iPhone and Polhemus Liberty, the sensor of the Polhemus Liberty was mounted on a Plexiglas holder which was attached to the strap that was holding the iPhone in place against the forehead (Fig. 2). The transmitter of the Polhemus Liberty was placed horizontally, squared, and centered in front of the subject, leveled with the sternum.

Before each measurement, the two instruments were calibrated simultaneously by the two raters. After 3 s, the Polhemus Liberty and the iPhone simultaneously registered the aCROM.

2.6.4. Interrater reliability

To assess interrater reliability, both raters performed the test sequence at an interval of 15 min, under the same conditions and in the same room. Based on clinical experience, we assumed this time frame to be long enough to allow for adequate recovery following cervical range of motion testing, and short enough to prevent changes in mobility and major variation due to rater or patient-related factors.

2.7. Sample size

To detect a minimal intraclass correlation coefficient (ICC) of 0.75 ($\alpha = 0.05$, $1-\beta = 0.80$), with an expected ICC of 0.9, and two measurements per participant with a 10% drop-out rate, a total of 26 participants was required for the reliability study (Walter et al., 1998). For the validity study, an equal sample was considered adequate, based on previous studies evaluating validity of cervical range of motion

Table 1
Preliminary validation on wooden phantom (frame): differences between Polhemus Liberty and iPhone.

	Difference Polhemus - iPhone Flexion-extension (x axis)		Difference Polhemus - iPhone Rotation (y axis)		Difference Polhemus - iPhone Lateral bending (z axis)	
	MD (95%CI)	LoA	MD (95%CI)	LoA	MD (95%CI)	LoA
Trial 1 + 2 + 3 (deg)	0.06 (-0.36 – 0.16)	-0.46 – 0.58	0.48 (0.37 – 0.59)	-0.12 – 1.08	0.10 (-0.01 – 0.21)	-0.49 – 0.69
Trial 1 (deg)	-0.05 (-0.15 – 0.04)	-0.31 – 0.20	0.23 (0.13 – 0.34)	-0.06 – 0.52	0.03 (-0.13 – 0.20)	-0.42 – 0.49
Trial 2 (deg)	0.02 (-0.15 – 0.19)	-0.46 – 0.49	0.46 (0.30 – 0.62)	0.01–0.91	0.09 (-0.12 – 0.30)	-0.50 – 0.67
Trial 3 (deg)	0.22 (-0.00 – 0.45)	-0.41 – 0.85	0.75 (0.56 – 0.94)	0.23–1.27	0.17 (-0.10 – 0.43)	-0.56 – 0.90

Deg = degrees; MD = mean difference; 95%CI = 95% confidence interval; LoA = limits of agreement.

measures (Assink et al., 2008; Guidetti et al., 2016; Quek et al., 2014; Tousignant-Laflamme et al., 2013).

2.8. Statistical analysis

Statistical analyses were performed using SPSS version 24.0 (SPSS Inc., Chicago, Illinois, USA). For both the validity study and the reliability study, the mean scores of the three repetitions were calculated and used for data analysis.

2.8.1. Concurrent validity

Validity was estimated using the intraclass correlation coefficient, two-way mixed effects model (ICC 3,1) to measure the association between scores from the two devices, and by calculating the limits of agreement (LoA = $1.96 \times \text{SD}$) (Bland and Altman, 1986; de Vet et al., 2011). ICCs of 0.75 and higher were considered to be acceptable for a useful measurement instrument (de Vet et al., 2011).

Coupled motion was defined as an association of one cervical motion (rotation or lateral bending) about an axis with another motion about a second axis. 3D coupling motion analysis consisted of two analysis methods (Cattrysse et al., 2012, 2008, 2006). First, to characterize the relationship between the main motion (e.g. rotation) and the coupled motion (e.g. lateral bending), cross-correlation coefficients between the main motion and the coupled motion were calculated for both instruments. A positive cross-correlation expresses an ipsilateral coupling pattern, a negative correlation a contralateral one (Cattrysse et al., 2012, 2008, 2006). Second, the ratio between the main and coupled motion was compared between both devices. The ratio expresses the amount of coupled motion relative to the main motion component (Cattrysse et al., 2012, 2008, 2006). The ratio was calculated as the standard deviation of the main motion divided by the standard deviation of the coupled motion (ratio = $\text{SD main motion} / \text{SD coupled motion}$) (Cattrysse et al., 2012, 2008, 2006). Data on cross correlation and ratio were visually checked for normality with histograms, Q-Q plots and boxplots, and tested with the Shapiro-Wilk test.

Differences in cross-correlation and ratio between the Polhemus Liberty and the iPhone were analyzed using the Wilcoxon signed-ranks test in case of not normally distributed data. The level of significance for all tests was set as $\alpha = 0.05$.

2.8.2. Interrater reliability

In order to estimate interrater reliability, intraclass correlation coefficients, for two-way random effects models (ICC 2,1), were calculated. Since in clinical practice different physical therapists perform the measurements, systematic differences between raters were considered to be part of the measurement error. ICC values of at least 0.70 were considered to be a minimal requirement for interrater reliability (de Vet et al., 2011).

Further, limits of agreement (LoA) were calculated to determine the absolute measurement error between the two raters (Bland and Altman, 1986; de Vet et al., 2011).

3. Results

3.1. Patient characteristics

Thirty patients (63.3% female) with a mean age of 53.4 years (SD 12.2) participated in the validity study, and the reliability sample consisted of 26 patients (73.1% female) with a mean age of 45.2 years (SD 15.3) (Table 2). Some observations were not available for all subjects because of technical iPhone and/or Polhemus Liberty errors (Tables 3–5).

Table 2
Characteristics of the patients.

Characteristics	Validity study (n = 30)	Reliability study (n = 26)
Age (years) (mean, SD)	53.4 ± 12.2	45.2 ± 15.3
Female	19 (63.3%)	19 (73.1%)
BMI (mean, SD)	25.6 ± 3.6	24.7 ± 3.7
Duration of symptoms		
0–3 months	6 (20.0%)	7 (26.9%)
3–6 months	6 (20.0%)	3 (11.5%)
6–12 months	3 (10.0%)	4 (15.4%)
> 1 year	15 (50.0%)	12 (46.2%)
NRS (median, range)		
Average	5.5 (0–8)	5.0 (2–8)
Minimal	2.5 (0–6)	3.0 (0–7)
Maximal	8.0 (1–10)	8.0 (3–9)
NDI score (median, range)	9.5 (1–26)	11.4 (3–27)

SD = standard deviation; BMI = Body Mass Index; NRS = Numeric Rating Scale (ranging from 0 to 10 points. Maximum pain = 10 points); NDI = Neck Disability Index (10 items ranging from 0 to 5 points. Maximal disability = 50 points).

3.2. Concurrent validity

3.2.1. Planar movements

Mean differences of half cycle movements between the iPhone and Polhemus Liberty ranged from 0.7° (SD 1.4°) for left lateral bending to 4.3° (SD 1.9°) for right rotation. For full cycle movements, mean differences ranged from 1.5° (SD 2.4°) for lateral bending to 8.4° (SD 2.9°) for rotation (Table 3).

All ICC values exceeded 0.90, indicating excellent validity. The limits of agreement ranged from -1.85° to 3.43° for right lateral bending to 2.70° – 14.14° for total rotation (Table 3), indicating that the difference between the measurements are expected to vary within this range.

3.2.2. 3D coupling motions

For both rotation and lateral bending as main motions, there were no significant differences in cross-correlation coefficients between the iPhone and the Polhemus Liberty (Table 4). In addition, for both rotation as well as for lateral bending as the primary movement no significant differences were found for the ratio analyses between both devices (Table 4).

3.3. Interrater reliability

ICCs for interrater reliability were 0.90 (95%CI 0.78–0.95) for flexion-extension, 0.92 (95%CI 0.82–0.97) for rotation, and 0.96 (95%CI 0.90–0.98) for lateral bending (Table 5). The limits of agreement, indicating the measurement error between the two independent raters, were -11.97° to 15.19° for flexion-extension, -10.06° – 13.82° for rotation, and -10.95° to 9.93° for lateral bending (Table 5).

4. Discussion

This study demonstrates high concurrent validity and interrater reliability of a new iPhone application to measure planar aCROM and associated coupling motions in patients with neck pain. Based on the observed ICC values (≥ 0.90) for concurrent validity and interrater reliability, the iPhone application appears to be a suitable instrument for discriminative purposes in the clinical decision making process in neck pain patients. However, 95% LoA between the measurement devices were wide implying relatively large absolute measurement error.

There are no comparable smartphone studies assessing the aCROM in neck pain patients. Previous smartphone studies assessing aCROM in healthy subjects found moderate to good concurrent validity for flexion-extension and lateral bending, and poor to moderate validity for

Table 3

Validity of the iPhone compared to the Polhemus Liberty (planar movements and mean scores on both devices, expressed in degrees).

	iPhone Mean (SD) (deg)	Polhemus Mean (SD) (deg)	MD (SD) (deg)	ICC (95%CI)	LoA (deg)
<i>Total Flexion-extension (n = 27)</i>	100.7 (15.1)	96.5 (14.0)	4.1 (2.4)	0.95 (0.15–0.99)	–0.62–8.82
Flexion (n = 27)	49.5 (8.5)	46.2 (7.9)	3.3 (1.5)	0.91 (–0.01–0.98)	–0.21–6.35
Extension (n = 27)	51.1 (12.4)	50.3 (11.8)	0.8 (1.7)	0.99 (0.97–0.99)	–2.56–4.15
<i>Total Rotation (n = 29)</i>	128.9 (22.4)	120.5 (20.6)	8.4 (2.9)	0.92 (–0.01–0.98)	2.70–14.14
Right (n = 29)	63.7 (10.8)	59.4 (9.8)	4.3 (1.9)	0.91 (–0.01–0.98)	0.58–8.08
Left (n = 29)	65.2 (14.1)	61.1 (13.1)	4.1 (1.8)	0.95 (0.04–0.99)	0.60–7.61
<i>Total lateral bending (n = 30)</i>	61.4 (18.2)	59.9 (16.7)	1.5 (2.4)	0.99 (0.96–0.99)	–3.24–6.32
Right (n = 30)	29.6 (8.8)	28.9 (8.2)	0.8 (1.3)	0.98 (0.95–0.99)	–1.85–3.43
Left (n = 30)	31.7 (10.1)	31.0 (9.4)	0.7 (1.4)	0.99 (0.97–0.99)	–2.11–3.61

SD = standard deviation; MD = Mean difference; deg = degrees; ICC = intraclass correlation coefficient; 95%CI = 95% confidence interval; LoA = Limits of agreement.

Table 4

Three dimensional range of motion: cross-correlation and ratio.

	Polhemus	iPhone	p value
	Mean (SD)	Mean (SD)	
<i>Rotation (n = 29)</i>			
Cross correlation	0.55 (0.44)	0.57 (0.46)	.230
Ratio	12.29 (5.05)	11.38 (7.11)	.214
<i>Lateral bending (n = 30)</i>			
Cross correlation	–0.91 (0.19)	–0.91 (0.14)	.271
Ratio	2.03 (1.19)	2.10 (1.44)	.629

SD = standard deviation.

Table 5

Interrater reliability and mean scores of rater 1 and 2.

	Rater 1 Mean (SD) (deg)	Rater 2 Mean (SD) (deg)	ICC (95%CI)	LoA (deg)
Flexion- extension (n = 24)	101.88 (14.47)	101.49 (16.19)	0.90 (0.78–0.95)	–11.97–15.19
Rotation (n = 21)	117.09 (21.87)	115.08 (18.74)	0.96 (0.90–0.98)	–10.06–13.82
Lateral bending (n = 23)	54.73 (12.03)	54.34 (13.83)	0.92 (0.82–0.97)	–10.95–9.93

SD = standard deviation; deg = degrees; 95%CI = 95% confidence interval; LoA = Limits of agreement.

rotation (Quek et al., 2014; Tousignant-Laflamme et al., 2013). Another study comparing an iPhone 5c with an inclinometer in healthy subjects demonstrated excellent validity and reliability, with ICCs above 0.99 and 95% LoA of 1.5–3.0° (Guidetti et al., 2016).

Results of previous studies that assessed concurrent validity and reliability of iPhone applications for the measurement range of motion of joints other than the cervical spine, are in line with our results. Good to excellent validity and reliability were reported for lumbar spine, shoulder, ankle, and knee range of motion measurements (Jones et al., 2014; Kolber et al., 2013; Mitchell et al., 2014; Pourahmadi et al., 2016; Romero Morales et al., 2017), which strengthens legitimacy of the use of iPhone applications for range of motion measurement in physical therapy.

There have been no studies with smartphones being used for the examination of coupled motions. Analysis of the results of our preliminary validity study shows only small, negligible differences (ranging from 0.02° to 0.75°) around all three axes between the iPhone and the Polhemus Liberty, which supports the conclusion that the iPhone adequately registers data on all axes.

We found substantial differences between the results of the

measurements performed on patients compared to those from our preliminary validation on the wooden frame. A possible reason for this is that there may have been a magnetic field interference between the MEMS-sensors and the Polhemus Liberty during the measurements on patients. Sensors of the devices were placed further apart on the wooden frame than they were on patients. MEMS sensors are very sensitive to electromagnetic fields (Roetenberg et al., 2005). Influence of electromagnetic fields on the results of smartphone measurements have been reported before (Quek et al., 2014). A study in which there was no such influence showed noticeably higher validity and smaller measurement errors (Guidetti et al., 2016). Another reason for the discrepancies might be that, despite our fixation, movement of the sensor of the Polhemus Liberty relative to the iPhone and head was still possible, which could also have negatively influenced intertrial variability. As a consequence, the results of the present study could have underestimated validity and reliability of the iPhone application.

It is important for clinicians to appreciate the width of the 95% LoA between the measurement devices. The estimated 95% LoA indicate that differences up to 2.6° for lateral flexion and up to 5.7° for total rotation may exist when using these devices interchangeably. Absolute measurement errors indicate that change scores should exceed 10.4° for total lateral bending to 13.6° for total flexion-extension to detect changes that are not due to intertrial variability or to measurement error. This magnitude of measurement error is reported at least as large in other studies that assessed validity and reliability of aCROM measures (Assink et al., 2005; Fletcher and Bandy, 2008; Jordan, 2000; Quek et al., 2014). In order to reduce the measurement error in practice, it is recommended to make sure that raters are well trained, measurements are taken under controlled conditions and the average of multiple measurements is used to determine aCROM (Streiner and Norman, 2008).

It is important that the measurement error found between raters is smaller than the minimal clinically important difference in order to discriminate in scores between neck pain patients and those without neck pain. A recent study showed that at group level differences between patients with neck pain and healthy persons are larger, with a minimum of 7.0° for half cycle movements (lateral bending) and 16.7° for full cycle movements (total lateral bending) (Stenneberg et al., 2016). Therefore, the clinically important difference seems to be beyond the measurement error found in the present study. We suggest further investigation of the clinically important difference in neck pain patients compared to individuals without neck pain to consider the clinical relevance of aCROM measurements.

4.1. Limitations

High ICCs could be explained by high variability in aCROM between patients (de Vet et al., 2006). However, the variability in aCROM among patients of our sample was comparable to that of other clinically representative samples (Stenneberg et al., 2016). We therefore assume

our ICC values to be a valid estimate of reliability.

Further, intrarater reliability was not assessed in this study. In clinical settings, therapists evaluate aCROM over time which is why an estimation of the intrarater reliability could be required as well. However, examining intrarater reliability seems to be of less importance when an instrument shows good interrater reliability (Streiner and Norman, 2008), as intrarater reliability usually tends to be higher than interrater reliability (Streiner and Norman, 2008).

Last, the technology used in smartphones innovates rapidly. New smartphones contain more advanced sensors and software. It is expected that newer and better technology will lead to more accurate measurements. The clinimetric properties of these new smartphones will need continuous validity and reliability evaluation.

5. Conclusions

Our results suggest that the new iPhone application is a useful diagnostic tool for the measurement of planar aCROM and associated coupling motions in patients with neck pain. As such it fulfills the need for a valid, reliable, and feasible instrument for measuring aCROM in clinical practice and research. Therapists and researchers should, however, consider measurement error when interpreting scores.

Conflicts of interest

All authors declare that there are no conflicts of interest.

Acknowledgement

The authors would like to thank Simon Busstra for the development of the iPhone-application.

References

- Antonaci, F., Ghirmai, S., Bono, G., Nappi, G., 2000. Current methods for cervical spine movement evaluation: a review. *Clin. Exp. Rheumatol.* 18.
- Assink, N., Bergman, G.J.D., Knoester, B., Winters, J.C., Dijkstra, P.U., Postema, K., 2005. Interobserver reliability of neck-mobility measurement by means of the flock-of-birds electromagnetic tracking system. *J. Manip. Physiol. Ther.* 28, 408–413. <http://dx.doi.org/10.1016/j.jmpt.2005.06.009>.
- Assink, N., Bergman, G.J.D., Knoester, B., Winters, J.C., Dijkstra, P.U., 2008. Assessment of the cervical range of motion over time, differences between results of the Flock of Birds and the EDI-320: a comparison between an electromagnetic tracking system and an electronic inclinometer. *Man. Ther.* 13, 450–455. <http://dx.doi.org/10.1016/j.math.2007.05.012>.
- Bier, J., Scholten-Peeters, G., Staal, J., Pool, J., van Tulder, M., Beekman, E., et al., 2016. KNGF-richtlijn Nekpijn.
- Bland, J.M., Altman, D.G., 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 327, 307–310. [http://dx.doi.org/10.1016/S0140-6736\(86\)90837-8](http://dx.doi.org/10.1016/S0140-6736(86)90837-8).
- Bogduk, N., Mercer, S., 2000. Biomechanics of the cervical spine. I: normal kinematics. *Clin Biomech (Bristol, Avon)* 15, 633–648. [http://dx.doi.org/10.1016/S0268-0033\(00\)00034-6](http://dx.doi.org/10.1016/S0268-0033(00)00034-6).
- Cattrysse, E., Baeyens, J.P., Kool, P., Clarys, J.P., Van Roy, P., 2006. In vitro 3D-artrokinematic analysis of axial rotation and lateral bending mobilization of the upper cervical spine. In: 9th Int Conf 3D-human Mov Anal Val.
- Cattrysse, E., Baeyens, J.P., Kool, P., Clarys, J.P., Van Roy, P., 2008. Does manual mobilization influence motion coupling patterns in the atlanto-axial joint? *J. Electromyogr. Kinesiol.* 18, 838–848. <http://dx.doi.org/10.1016/j.jelekin.2007.02.015>.
- Cattrysse, E., Moens, M., Schailée, E., D'Haens, J., Van Roy, P., 2012. Changed cervical kinematics after fusion surgery. *Eur. Spine J.* 21, 1353–1359. <http://dx.doi.org/10.1007/s00586-011-2112-z>.
- Childs, J.D., Cleland, J.A., Elliott, J.M., Teyhen, D.S., Wainner, R.S., Whitman, J.M., et al., 2008. Neck pain. *J. Orthop. Sports Phys. Ther.* 38, A1–A34. <http://dx.doi.org/10.2519/jospt.2008.0303>.
- Cook, C., Hegedus, E., Showalter, C., Sizer, P.S., 2006. Coupling behavior of the cervical spine: a systematic review of the literature. *J. Manip. Physiol. Ther.* 29, 570–575. <http://dx.doi.org/10.1016/j.jmpt.2006.06.020>.
- Feipel, V., Rondelet, B., Le Pallec, J., Rooze, M., 1999. Normal global motion of the cervical spine. *Clin. Biomech.* 14, 462–470. [http://dx.doi.org/10.1016/S0268-0033\(98\)90098-5](http://dx.doi.org/10.1016/S0268-0033(98)90098-5).
- Ferriero, G., Sartorio, F., Foti, C., Primavera, D., Brigatti, E., Vercelli, S., 2011. Reliability of a new application for smartphones (DrGoniometer) for elbow angle measurement. *PM R* 3, 1153–1154. <http://dx.doi.org/10.1016/j.pmrj.2011.05.014>.
- Fletcher, J.P., Bandy, W.D., 2008. Intrarater reliability of CROM measurement of cervical spine active range of motion in persons with and without neck pain. *J. Orthop. Sports Phys. Ther.* 38, 640–645. <http://dx.doi.org/10.2519/jospt.2008.2680>.
- Guidetti, L., Placentino, U., Baldari, C., 2016. Reliability and criterion validity of the smartphone inclinometer application to quantify cervical spine mobility. *Clin Spine Surg* 0, 1–8. <http://dx.doi.org/10.1097/BSD.0000000000000364>.
- Guo, L.Y., Lee, S.Y., Lin, C.F., Yang, C.H., Hou, Y.Y., Wu, W.L., et al., 2012. Three-dimensional characteristics of neck movements in subjects with mechanical neck disorder. *J. Back Musculoskelet. Rehabil.* 25, 47–53. <http://dx.doi.org/10.3233/BMR-2012-0309>.
- Guzman, J., Hurwitz, E.L., Carroll, L.J., Haldeman, S., Côté, P., Carragee, E.J., et al., 2009. A new conceptual model of neck pain: linking onset, course, and care: the bone and joint decade 2000–2010 task force on neck pain and its associated disorders. *J. Manip. Physiol. Ther.* 32, S17–S28. <http://dx.doi.org/10.1016/j.jmpt.2008.11.007>.
- Jasiewicz, J.M., Treleaven, J., Condie, P., Jull, G., 2007. Wireless orientation sensors: their suitability to measure head movement for neck pain assessment. *Man. Ther.* 12, 380–385. <http://dx.doi.org/10.1016/j.math.2006.07.005>.
- Jones, A., Sealey, R., Crowe, M., Gordon, S., 2014. Concurrent validity and reliability of the Simple Goniometer iPhone app compared with the Universal Goniometer. *Physiother. Theor. Pract.* 3985, 1–5. <http://dx.doi.org/10.3109/09593985.2014.900835>.
- Jordan, K., 2000. Review of the literature assessment of published reliability studies for cervical spine range-of-motion measurement tools. *J. Manip. Physiol. Ther.* 23, 180–195. <http://dx.doi.org/10.1067/mmt.2000.105114>.
- Kaliarntas, K., Ugolue, U., Riches, P., Rowe, P., 2009. Concurrent validity and test-retest reliability of the Polhemus Liberty for the measurement of spinal range. In: *Int Soc Biomech XXII World Congr 2009-07-05-2009-07-09*, Cape Town, South Africa.
- Kolber, M.J., Pizzini, M., Robinson, A., Yanez, D., Hanney, W.J., 2013. The reliability and concurrent validity of measurements used to quantify lumbar spine mobility: an analysis of an iPhone® application and gravity based inclinometry. *Int J Sports Phys Ther* 8, 129–137.
- de Koning, C.H.P., van den Heuvel, S.P., Staal, J.B., Smits-Engelsman, B.C.M., Hendriks, E.J.M., 2008. Clinimetric evaluation of active range of motion measures in patients with non-specific neck pain: a systematic review. *Eur. Spine J.* 17, 905–921. <http://dx.doi.org/10.1007/s00586-008-0656-3>.
- Kottner, J., Audige, L., Brorson, S., Donner, A., Gajewski, B.J., Hróbjartsson, A., et al., 2011. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *Int. J. Nurs. Stud.* 48, 661–671. <http://dx.doi.org/10.1016/j.ijnurstu.2011.01.016>.
- Law, E.Y.H., Chiu, T.T.W., 2013. Measurement of Cervical Range of Motion (CROM) by electronic CROM goniometer: a test of reliability and validity. *J. Back Musculoskelet. Rehabil.* 26, 141–148. <http://dx.doi.org/10.3233/BMR-2012-00358>.
- Luinge, H., 1999. Estimating orientation with gyroscopes and accelerometers. *Technol. Health Care* 7, 455–459.
- MacDermid, J.C., Walton, D.M., Avery, S., Blanchard, A., Etruw, E., McAlpine, C., et al., 2009. Measurement properties of the neck disability index: a systematic review. *J. Orthop. Sports Phys. Ther.* 39, 400–417. <http://dx.doi.org/10.2519/jospt.2009.2930>.
- Mitchell, K., Gutierrez, S.B., Sutton, S., Morton, S., Morgenthaler, A., 2014. Reliability and validity of goniometric iPhone applications for the assessment of active shoulder external rotation. *Physiother. Theor. Pract.* 3985, 521–525. <http://dx.doi.org/10.3109/09593985.2014.900593>.
- Nafis, C., Jensen, V., Beauregard, L., Anderson, P., 2006. Method for estimating dynamic EM tracking accuracy of Surgical Navigation Tools. *Proc SPIE Med Imag Vis Image-Guided Procd Model* 6141, 152–167. <http://dx.doi.org/10.1117/12.653448>.
- Ockendon, M., Gilbert, R., 2012. Validation of a novel smartphone accelerometer-based knee goniometer. *J. Knee Surg.* 25, 341–346. <http://dx.doi.org/10.1055/s-0031-1299669>.
- Pool, J.J.M., Ostelo, R.W.J.G., Hoving, J.L., Bouter, L.M., de Vet, H.C.W., 2007. Minimal clinically important change of the neck disability index and the numerical rating Scale for patients with neck pain. *Spine (Phila Pa 1976)* 32, 3047–3051. <http://dx.doi.org/10.1097/BRS.0b013e31815c7f5b>.
- Pourahmadi, M.R., Taghipour, M., Jannati, E., Mohseni-Bandpei, M.A., Ebrahimi Takamjani, I., Rajabzadeh, F., 2016. Reliability and validity of an iPhone® application for the measurement of lumbar spine flexion and extension range of motion. *PeerJ* 4, e2355. <http://dx.doi.org/10.7717/peerj.2355>.
- Quek, J., Brauer, S.G., Treleaven, J., Pua, Y.-H., Mentiply, B., Clark, R.A., 2014. Validity and intra-rater reliability of an android phone application to measure cervical range-of-motion. *J. NeuroEng. Rehabil.* 11, 65. <http://dx.doi.org/10.1186/1743-0003-11-65>.
- Roetenberg, D., Luinge, H.J., Baten, C.T.M., Veltink, P.H., 2005. Compensation of magnetic disturbances improves inertial and magnetic sensing of human body segment orientation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 13, 395–405. <http://dx.doi.org/10.1109/TNSRE.2005.847353>.
- Romero Morales, C., Calvo Lobo, C., Rodríguez Sanz, D., Sanz Corbalán, I., Ruiz Ruiz, B.B., López López, D., 2017. The concurrent validity and reliability of the Leg Motion system for measuring ankle dorsiflexion range of motion in older adults. *PeerJ* 5, e2820. <http://dx.doi.org/10.7717/peerj.2820>.
- Schellenghouth, J.M., Verhagen, A.P., Heymans, M.W., Koes, B.W., De Vet, H.C., Terwee, C.B., 2012. Measurement properties of disease-specific questionnaires in patients with neck pain: a systematic review. *Qual. Life Res.* 21, 659–670. <http://dx.doi.org/10.1007/s11136-011-9965-9>.
- Shin, S.H., Ro, D.H., Lee, O.S., Oh, J.H., Kim, S.H., 2012. Within-day reliability of shoulder range of motion measurement with a smartphone. *Man. Ther.* 17, 298–304. <http://dx.doi.org/10.1016/j.math.2012.02.010>.
- Stenneberg, M.S., Rood, M., de Bie, R., Schmitt, M.A., Cattrysse, E., Scholten-Peeters,

- G.G., 2017. To what degree Does active cervical range of motion differ between patients with neck pain, patients with whiplash, and those without neck Pain? A systematic review and meta-analysis. *Arch. Phys. Med. Rehabil.* 98, 1407–1434. <http://dx.doi.org/10.1016/j.apmr.2016.10.003>.
- Streiner, D., Norman, G., 2008. *Health Measurement Scales. A Practical Guide to Their Development and Use*. University Press, Oxford. <http://dx.doi.org/10.1378/chest.96.5.1161>.
- Theobald, P.S., Jones, M.D., Williams, J.M., 2012. Do inertial sensors represent a viable method to reliably measure cervical spine range of motion? *Man. Ther.* 17, 92–96. <http://dx.doi.org/10.1016/j.math.2011.06.007>.
- Tousignant-Laflamme, Y., Boutin, N., Dion, A.M., Vallée, C.-A., 2013. Reliability and criterion validity of two applications of the iPhone™ to measure cervical range of motion in healthy participants. *J. NeuroEng. Rehabil.* 10, 1. <http://dx.doi.org/10.1186/1743-0003-10-69>.
- Vernon, H., Mior, S., 1991. The Neck Disability Index: a study of reliability and validity. *J. Manip. Physiol. Ther.* 409–415.
- de Vet, H.C.W., Terwee, C.B., Knol, D.L., Bouter, L.M., 2006. When to use agreement versus reliability measures. *J. Clin. Epidemiol.* 59, 1033–1039. <http://dx.doi.org/10.1016/j.jclinepi.2005.10.015>.
- de Vet, H.C.W., Terwee, C.B., Mokkink, L.B., Knol, D.L., 2011. *Measurement in Medicine: a Practical Guide*. Cambridge University Press.
- Walter, S.D., Eliasziw, M., Donner, A., 1998. Sample size and optimal designs for reliability studies. *Stat. Med.* 17, 101–110. [http://dx.doi.org/10.1002/\(SICI\)1097-0258\(19980115\)17:1<101::AID-SIM727>3.0.CO;2-E](http://dx.doi.org/10.1002/(SICI)1097-0258(19980115)17:1<101::AID-SIM727>3.0.CO;2-E).
- Williams, M.A., McCarthy, C.J., Chorti, A., Cooke, M.W., Gates, S., 2010. A systematic review of reliability and validity studies of methods for measuring active andPassive cervical range of motion. *J. Manip. Physiol. Ther.* 33, 138–155. <http://dx.doi.org/10.1016/j.jmpt.2009.12.009>.
- Woodhouse, A., Vasseljen, O., 2008. Altered motor control patterns in whiplash and chronic neck pain. *BMC Musculoskel. Disord.* 9, 90. <http://dx.doi.org/10.1186/1471-2474-9-90>.